AD-A260 526

KEEP THIS COPY FOR REPRODUCTION PURPOSES

JMENTATION PAGE

Form Approved
OMB No. 0704-0188

(2)

...s estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, ...ting and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this ...cing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson ...d to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE 10-21-92	3. REPORT TYPE AND DATES COVERED Final Report - Jan '89 to Jan '92

4. TITLE AND SUBTITLE

Neural Net Architecture For Computing Structure From Motion

5. FUNDING NUMBERS

DALL03-88-K-0052

6. AUTHOR(S)

Josef Skrzypek

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Machine Perception Laboratory
Computer Science Department
UCLA, Los Angeles, CA 90024

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ARO 25324.10-MA

11. SUPPLEMENTARY NOTES

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

93-03114

8PG

13. ABSTRACT (*Maximum 200 words*)

Analysis of motion contributes to the image understanding tasks by disambiguating scene information whenever, the observer and/or objects in the scene are in motion. This proposal is focused on research and development of algorithms for automatic recalibration from sensory to egocentric coordinates during egomotion.

14. SUBJECT TERMS

15. NUMBER OF PAGES	6
16. PRICE CODE	

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NEURAL NET ARCHITECTURE FOR COMPUTING STRUCTURE FROM MOTION

Josef Skrzypek

10-21-92

U.S. ARMY RESEARCH OFFICE

DALL03-88-K-0052

Machine Perception Laboratory
Computer Science Department
University of California
Los Angeles CA 90024

Approved for public release

Distribution unlimited

Neural net architecture for computing structure from motion

Josef Skrzypek

Machine Perception Laboratory
Computer Science Department
University of California
Los Angeles, CA 90024

November 23, 1992

1 ABSTRACT

Analysis of motion contributes to the image understanding tasks by disambiguating scene information whenever, the observer and/or objects in the scene are in motion. This proposal is focused on research and development of algorithms for automatic recalibration from sensory to egocentric coordinates during egomotion.

2 Statement of the problem

2.1 Introduction

The computation of shape from motion cannot be done robustly by a system that passively interprets visual information. A system capable of operating in unconstrained natural environments must have the capability of modifying its own sensory input – through self-motion, manipulation of objects within the environment, or selective focusing of processing power onto limited regions of the environment.

Based on this hypothesis and substantial evidence from Neuroscience and Psychology, an architecture for active perceptual processing has been developed. This architecture includes tightly coupled afferent (sensory) and efferent (manipulatory) subsystems.

A common frame of reference is necessary for these afferent and efferent systems. Therefore, simulations have been performed on one part of the architecture which converts visual input from a sensory-array-based coordinate system to an egocentric coordinate system. This egocentric coordinate system is the same one which is used for selection and coordination of efferent, exploratory actions.

2.2 Methods

All simulations results were generated at the UCLA Machine Perception Laboratory using neural network simulation environment called SFINX (Structure and Function In Neural conneXions)[6,5,14,4], running on Ardent Titans and other UNIX machines. SFINX allows simulation of both "randomly" and regularly interconnected networks, where individual "neurons" can be simulated at almost any level of detail, from a set of difference equations to a simplified weighted-sum model. To accomplish this simulation, functions must be written that simulate individual neurons, with SFINX taking care of applying such functions to each neuron in the net. The functions have flexibility to define not only the operation of individual elements, but also the interconnection

between elements. SFINX includes facilities for displaying the internal state of elements as images. Additionally, the state may be saved to a file for further processing.

Using UCLA-SFINX, neural network architectures have been developed to perform simple foveation in specified directions by computing a transformation from retinocentric to egocentric coordinates. Thus the selective attention mechanism operates on the output of the foveation module in the egocentric coordinate system. It is composed of four two-dimensional layers of "neurons", each of which is modeled after visual cortical cells. These layers perform, in stages, a shifting of sensor array input to its proper location within an egocentric map.

Questions related to what is meant by shape and in what ways can that information be extracted from motion have not been directly addresed in this project. Open problems include: what is the process by which attention (as in the motion of the sensor array) is directed? How to segment local object motion from the background? How is localized motion related to determining spatial location of objects? How are localized motions grouped together to form potential targets for attention? What criteria are used to select a target? What guides the shift of attention? Answer to these questions is a necessary step towards the development of a machine architecture which can analyze the spatial relations between objects in the environment, and coordinate motor operations within the environment. Once the architecture can perform these operations, it will be possible to extract detailed object shape, independent of object position.

3 Summary of the results

Projections of a moving target on the sensory array undergo continuous shifts, especially during egomotion. Without minimal compensation, such as retinal to egocentric conversion, it would be difficult to realize a meaningful interpretation of the environment. We developed a model that can perform such a conversion, without resorting to learning schemes, and we showed that its structure is consistent with the known anatomy of the visual system. The model preattentively discounts changes in visual direction, allowing attentive processes to generate a stable and consistent internal representation of the surrounding world.

Our model does not require learning which has been suggested to account for target motion represented by the single spots of light[?]. In normal operation, the visual system operates on complex patterns of light. Our model works with arbitrary visual input, and its responses to more complex experimental stimuli can be compared with similar experiments performed on animals to validate or invalidate the model.

If we assume that the granularity of the egocentric map is such that it represents

space at the maximum visual acuity achievable by humans, at most five hierarchical stages of neurons would be required to perform the transform. There is no problem fitting such an architecture within the known neuroanatomy of visual areas.

Our model of the egocentric transform makes a low demand on the processing power and interconnectivity of neurons. Thus, the same neurons can be involved in simulataneous other processing, such as segmenting images based on motion. The result in real terms would be neurons which not only respond to eye position, but also *attributes* of the visual stimulation (for example, the overall optic flow).

Finally, with respect to the work of Treisman and others in attention, our model predicts that the "master map of locations", within which attention (both internal and motor) operates, is an egocentric map. It favors a model in which this map is placed *after* feature extraction modules.

A facility has been constructed to produce real-world motion sequences for input to the UCLA SFINX neural network simulation environment []. Output images can then be saved on videotape for replay in real-time. This facility includes computer-controlled video recorders, editors, and a video laser disk player, all controlled by the SFINX software.

4 List of publications

[19,21,20,2,7,3,9,14,10,12,13,1,8,16,17,18,15,11,?]

References

[1] R. Gleeson and J. Skrzypek. Boundaries as unpredictable discontinuities. In *Proceedings of SPIE Conference on Image Understanding and the Man-Machine Interface*, volume 1076, Los Angeles, California, January 1989.

[2] W. Lincoln and J. Skrzypek. Dynamics of clustering multiple back propagation networks. In *Proceedings of IEEE Intl. Conference on Systems, Man, and Cybernetics*, Los Angeles, California, October 1990.

[3] E. Mesrobian and J. Skrzypek. A tool for simulating neural models. In *Proceedings of the IEEE International Conference on Man, Systems, and Cybernetics*, Los Angeles, California, November 1990.

[4] E. Mesrobian and J. Skrzypek. A Software Environment for Studying Computational Neural Systems. *IEEE Transactions on Software Engineering*, 18(7):575–589, July 1992.

[5] E. Paik, D. Gungner, and J. Skrzypek. UCLA SFINX - A neural network simulation environment. In *Proceedings of the IEEE First Annual International Conference on Neural Networks*, volume 3, pages 367–376, San Diego, California, June 1987.

[6] E. Paik and J. Skrzypek. Ucla-sfinx - neural network simulation environment. Technical Report UCLA-MPL-TR 4-87, University of California Los Angeles Machine Perception Laboratory, April 1991.

[7] V. Rodrigues and J. Skrzypek. Categorizing visual stimuli: Specification of a neural network architecture. In *Proceedings of the IEEE Intl. Conference on Systems, Man, and Cybernetics*, Los Angeles, California, November 1990.

[8] J. Skrzypek. Lightness Constancy: Connectionist Architecture for Controlling Sensitivity. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(5):957–968, September/October 1990.

[9] J. Skrzypek. Neural specification of a general purpose vision system. In *Proceedings of ACNN'90, Australian Conference on Neural Networks*, page 160, Sydney, Australia, January 1990. Invited paper.

[10] J. Skrzypek. Feedback Synapse to Cone and Light Adaptation. In D Touretzky, editor, *Advances in Neural Information Processing Systems, NIPS-3*, pages 391–396, San Mateo, CA, 1991. Morgan Kaufman.

[11] J. Skrzypek. Neural network for image segmentation;: Illusory contours. In *Proceedings 2nd Government Neural network Application Workshop*, Huntsville, Alabama, 1991.

[12] J. Skrzypek. Light sensitivity in cones is affected by the feedback from horizontal cells. In F. Eeckman, editor, *Analysis and Modeling of Neural Systems*, pages 213–223. Kluwer Academic, 1992.

[13] J. Skrzypek and J. Hoffman. Visual Recognition of Script Characters and Neural Network Architectures. In E. Gelenbe, editor, *Neural Networks: Advances and Applications*, pages 109–144. Elsevier Science Publishers, North Holland, 1991.

[14] J. Skrzypek and E. Mesrobian. UCLA-SFINX — Simulating Structure and Function in Neural Connections. *Transactions of The Society for Computer Simulation*, 8(3):181–217, September 1991.

[15] J. Skrzypek and B. Ringer. Neural Network Models for Illusory Contour Perception. Technical Report UCLA-MPL-TR 91-12, Machine Perception Laboratory, University of California, Los Angeles, October 1991.

[16] J. Skrzypek, E. Tisdale, and K. Frankel. Neural networks for recognition of handwritten script characters: comparative analysis. In *Proceedings of Second Intl. Workshop on Frontiers in Handwriting Recognition*, Chateau de Bonas, France, September 1991.

[17] J. Skrzypek and G. Wu. Invariant Contrast Adaptation in the Primate Outer Plexiform Layer (OPL). Technical Report UCLA-MPL-TR 91-13, Machine Perception Laboratory, University of California, Los Angeles, November 1991.

[18] J. Skrzypek and G. Wu. Neither DoG nor LoG fits the Receptive Field of the Vertebrate Cone. In F. Eeckman, editor, *Analysis and Modeling of Neural systems II*. Kluwer Academic, 1992. in press.

[19] M. Stiber and J. Skrzypek. An architecture for visual direction constancy: Towards shape from motion. Technical Report UCLA-MPL-TR 90-1, University of California, Los Angeles Machine Perception Laboratory, April 1990.

[20] M. Stiber and J. Skrzypek. Peripheral Motion Triggered Attention. In *Proeedings of IEEE Intl. Conference on Systems, Man, and Cybernetics*, Los Angeles, CA, November 1990.

[21] M. Stiber and J. Skrzypek. An Architecture for Visual Direction Constancy: Towards Shape from Motion. *Journal of Artificial Neural Networks*, 1992. in press.

5 List of participating scientific personnel

Michael Stiber, Master Thesis 1990.
David Gungner, PhD student
Brian Ringer. Master Thesis 1992
Geurge Wu, Master Thesis 1992